数理解析研究所講究録 1462 巻 2006 年 202-210

Discrete Itô Formulas and Their Applications to Stochastic Numerics

赤堀 次郎
Jirô Akahori

立命館大学理工学部
Department of Mathematical Sciences
Ritsumeikan University

1-1-1, Nojihigashi, Kusatsu, Shiga, 525-8577, Japan

E-mail: akahori@se.ritsumei.ac.jp

2000 Mathematics Subject Classifications:60-02, 60-06, 65-C30

1 Introduction

This is a survey of the author's observations on the discrete-time analogues of Itô formulas. The observations presented here are summarized as follows.

1. As is well known, the standard Itô formula is based on the stochastic integral, which we do not need in discrete-time framework. To have an equality, we instead rely on the Fourier series expansion. Detailed explanations will be given in Section 2.

2. In a parallel way that the standard one describes the Kolmogorov equation for a given stochastic differential equation, the discrete Itô formula gives a finite difference equation for a given approximating equation (Euler-Maruyama, for example) of SDE. This observation leads to a computational framework of Monte-Carlo simulations of the finite difference scheme for partial differential equations in a high dimension. (Section 3.3)

3. The algebraic nature of the discrete Itô calculus fits well to the so-called *lattice framework* (for approximations) in mathematical finance. The Itô

formula reveals that it is nothing but a variant of Euler-Maruyama scheme. (Section 2.6 and Section 3.2)

2 Discrete Itô formulas

2.1 Fujita's Itô formula

Let us start with Fujita's Discrete Itô Formula (DIF for short) [3] for the simple random walk;

$$W_k = \tau_1 + \tau_2 + \cdots + \tau_k, \tag{2.1}$$

where $\{\tau_1, ..., \tau_t, ...\}$ is a Bernoulli sequence such that $\mathsf{P}(\tau_k = \pm 1) = 1/2$. By a simple algebra, we have

$$F(\tau_k) = \frac{\{F(+1) - F(-1)\}}{2}\tau_k + \frac{\{F(+1) + F(-1)\}}{2} \tag{2.2}$$

for every $F : \{-1, +1\} \to \mathbf{R}$. A DIF for $f : \mathbf{Z} \to \mathbf{R}$ is obtained by regarding $f(W_{t+1}) = f(W_t + \tau_t)$ as a function on $\{-1, +1\}$ and applying (2.2) to it:

$$\begin{aligned} f(W_{t+1}) - f(X_t) &= f(W_t + \tau_{t+1}) - f(W_t) \\ &= \frac{f(W_t + 1) - f(W_t - 1)}{2}\tau_{t+1} + \frac{f(W_t + 1) + f(W_t - 1)}{2} - f(W_t). \end{aligned}$$

The starting point is to regard (2.2) as Fourier expansion of F with respect to the orthonormal basis $\{1, \tau_k\}$.

2.2 The first generalization

Let ξ be a real valued random variable with $\mathbf{E}[\xi] = 0$ and $\mathrm{Var}[\xi] = \mathbf{E}[\xi^2] < \infty$. Let ν be its law and choose an orthonormal basis $\{H_n : n \in \mathbf{N}\}$ of the Hilbert space $L^2(\mathbf{R}, \nu)$ by expanding $H_0 \equiv 1, H_1(x) = x/\sqrt{\mathrm{Var}[\xi]}$. Then the DIF for the random walk (the sum of independent copies of ξ)

$$W_{t_k} = \xi_1 + \xi_2 + \cdots + \xi_k, \quad (k \in \mathbf{Z})$$

would be the following orthogonal expansion of $f(W_{t_{k+1}} + \ \cdot\ ) - f(W_{t_k})$.

$$\begin{aligned} f(W_{t_{k+1}}) - f(W_{t_k}) &= f(W_{t_k} + \xi_{k+1}) - f(W_{t_k}) \\ &= \frac{1}{\mathbf{E}[|\xi|^2]}\left(\int f(W_{t_k} + x) x\, \nu(dx)\right)(W_{t_{k+1}} - W_{t_k}) \\ &\quad + \frac{1}{t_{k+1} - t_k}\left(\int \{f(W_{t_k} + x) - f(W_{t_k})\}\, \nu(dx)\right)(t_{k+1} - t_k) \\ &\quad + \sum_{n=2}^{\infty}\left(\int f(W_{t_{k+1}} + x) H_n(x)\, \nu(dx)\right) H_n(\xi_{k+1}). \end{aligned} \tag{2.3}$$

Here f is a bounded measurable function.

When ξ is Gaussian, $\{H_n\}$ could be the Hermite polynomials (up to constants). Further, if $\mathbf{E}[\xi_k^2] = t_k - t_{k-1}$, then W can be a Brownian motion.

2.3 A multi-dimensional extension

The DIF for a multi-dimensional random walk $\mathbf{W} = (W^1, ..., W^n)$, where $W^j_{t_k} - W^j_{t_{k-1}} = \xi^j_k$, can be obtained through the expansion in the tensor product $\otimes_j L^2(\nu_j)$. Here the law of ξ^j is denoted by ν_j. Letting ν_0 be a trivial measure, we get a DIF of $(t, \mathbf{W})$ as

$$
\begin{aligned}
&f(t_{k+1}, \mathbf{W}_{t_{k+1}}) - f(t_k, \mathbf{W}_{t_k}) \\
&= \sum_j \frac{1}{\mathbf{E}[|\xi^j|^2]} \left(\int f(t_{k+1}, W^1_{t_k} + x_1, ..., W^j_{t_k} + x_j, ...) x_j \, \nu_j(dx_j) \right) (W^j_{t_{k+1}} - W^j_{t_k}) \\
&\quad + \frac{1}{t_{k+1} - t_k} \left(\int \{ f(t_{k+1}, \mathbf{W}_{t_k} + \mathbf{x}) - f(t_k, \mathbf{W}_{t_k}) \} \nu_1 \otimes \cdots \otimes \nu_n(d\mathbf{x}) \right) (t_{k+1} - t_k) \\
&\quad + \sum_{l_1 + \cdots + l_n \geq 2} \left(\int f(t_{k+1}, \mathbf{W}_{t_k} + \mathbf{x}) H^1_{l_1}(x_1) \cdots H^n_{l_n}(x_n) \nu_1(dx_1) \cdots \nu_n(dx_n) \right. \\
&\qquad\qquad\qquad \left. \cdot H^1_{l_1}(\xi^1_{t_{k+1}}) \cdots H^n_{l_n}(\xi^n_{t_{k+1}}) \right).
\end{aligned}
$$

Here f is a bounded measurable function on $\mathbf{R}^{n+1}$.

Now one sees that, in a sense, our DIF gives a discrete and conditioned chaos expansion (even when ξ's are not Gaussian). The 0-th and the 1st chaoses consist of the main terms and the higher order terms correspond to the *correction terms*, when one wants to get the standard Itô formula by letting $\Delta t := t_{k+1} - t_k$ to 0.

2.4 DIF for solutions to stochastic difference equations

Let $\mathbf{X}$ be the solution of a stochastic *difference* equation innovated by $\mathbf{W}$. That is,

$$
\mathbf{X}_{t_{k+1}} = \mathbf{X}_{t_k} + F(\mathbf{X}_{t_k}, t_{k+1} - t_k, \mathbf{W}_{t_{k+1}} - \mathbf{W}_{t_k}) \tag{2.4}
$$

for some vector field $F : \mathbf{R}^n \times \mathbf{R}^{n+1} \to \mathbf{R}^n$. The DIF for $(t, \mathbf{X})$ would be

$$
\begin{aligned}
& f(t_{k+1}, \mathbf{X}_{t_{k+1}}) - f(t_k, \mathbf{X}_{t_k}) \\
&= \sum_j \frac{1}{\mathbf{E}[|\xi^j|^2]} \left(\int f(t_{k+1}, \mathbf{X}_{t_k} + F(\mathbf{X}_{t_k}, \Delta t, \mathbf{x}))\, x_j\, \nu_j(dx_j) \right) (W^j_{t_{k+1}} - W^j_{t_k}) \\
&\quad + \frac{1}{t_{k+1} - t_k} \left(\int \{ f(t_{k+1}, \mathbf{X}_{t_k} + F(\mathbf{X}_{t_k}, \Delta_t, \mathbf{x})) - f(t_k, \mathbf{X}_{t_k}) \}\, \nu_1 \otimes \cdots \otimes \nu_n(d\mathbf{x}) \right) (t_{k+1} - t_k) \\
&\quad + \sum_{l_1 + \cdots + l_n \geq 2} \left(\int f(t_{k+1}, \mathbf{X}_{t_k} + F(\mathbf{X}_{t_k}, \Delta t, \mathbf{x})) H^1_{l_1}(x_1) \cdots H^n_{l_n}(x_n)\, \nu_1(dx_1) \cdots \nu_n(dx_n) \right. \\
&\qquad\qquad \left. \cdot H^1_{l_1}(\xi^1_{t_{k+1}}) \cdots H^n_{l_n}(\xi^n_{t_{k+1}}) \right).
\end{aligned}
\tag{2.5}
$$

When $F(\mathbf{x}, \Delta t, \mathbf{y})$ is affine in $\mathbf{y}$ as

$$
F(\mathbf{x}, \Delta t, \mathbf{y}) = \sigma(\mathbf{x})\mathbf{y}\sqrt{\Delta t} + \mu(\mathbf{x})\Delta t, \tag{2.6}
$$

where $\sigma : \mathbf{R}^n \to \mathbf{R}^n \otimes \mathbf{R}^n$ and $\mu : \mathbf{R}^n \to \mathbf{R}^n$, and when $\mathbf{W}$ is a Brownian motion, then (2.4) can be seen as an Euler-Maruyama scheme of a stochastic differential equation

$$
d\mathbf{X} = \sigma(\mathbf{X})\, d\mathbf{W} + \mu(\mathbf{X})\, dt.
$$

Still many classes, including higher order schemes and approximation schemes to SDE driven by Lévy processes, can be also written in the form of (2.4).

2.5 DIF for a class of weak approximation schemes

For a weak approximation scheme in a Brownian cases, we introduce another framework. If we define an n-dimensional random walk $\mathbf{W} = (W^1, ..., W^n)$ by $W^j_{t_k} - W^j_{t_{k-1}} = H_j(\xi_k)\sqrt{\Delta t}$, the (2.4) will work as a weak approximation scheme, based on the fact that

$$
\sqrt{\Delta t} \left(\sum_{t_k \leq t} H_1(\xi_{t_k}), \cdots, \sum_{t_k \leq t} H_n(\xi_{t_k}) \right), \; t \geq 0
$$

converges in law to the d-dimensional Wiener process due to the martingale central limit theorem (see, e.g. [2, Chapter 7]).

In this framework, the DIF for $(t, \mathbf{X})$ becomes

$$
\begin{aligned}
&f(t_{k+1}, \mathbf{X}_{t_{k+1}}) - f(t_k, \mathbf{X}_{t_k}) \\
&= \sum_{j=1}^{n} \frac{1}{\sqrt{\Delta t}} \left(\int f(t_{k+1}, \mathbf{X}_{t_k} + F(\mathbf{X}_{t_k}, \Delta t, \mathbf{H}(x)))\, H_j(x)\, \nu(dx) \right) (W^j_{t_{k+1}} - W^j_{t_k}) \\
&\quad + \frac{1}{\Delta t} \left(\int \{ f(t_{k+1}, \mathbf{X}_{t_k} + F(\mathbf{X}_{t_k}, \Delta t, \mathbf{H}(x))) - f(t_k, \mathbf{X}_{t_k}) \}\, \nu(dx) \right) (t_{k+1} - t_k) \\
&\quad + \sum_{l > n} \left(\int f(t_{k+1}, \mathbf{X}_{t_k} + F(\mathbf{X}_{t_k}, \Delta t, \mathbf{H}(x)))\, H_l(x)\, \nu(dx) \right) H_l(\xi_{t_{k+1}}),
\end{aligned}
\tag{2.7}
$$

where we have denoted $\mathbf{H}(x) = (H_1(x), \ldots, H_n(x))$.

2.6 DIF for complete markets

If $\sharp G = n+1$ in section 3.1, then $\{H_0, \ldots, H_n\}$ spans the whole space and the correction terms in (2.7) disappear. That is, the DIF becomes symbolically equivalent to the standard one. This is because the cardinality of the martingale basis of $\mathbf{W}$ is equal to the dimension of the state space of itself, as is the case with the standard Brownian motions.

From a perspective of mathematical finance, this property is closely related to *completeness* of the markets modeled by the stochastic process.

Roughly speaking, a market is said to be complete if every good has a unique price that excludes arbitrage opportunities. In many models in financial engineering the market is assumed to be complete to avoid discussing too much about the utilities/preferences of individuals.

Discrete-time complete market models are studied in [1] using a DIF.

2.7 Supplementary remarks for section 2

We remark that:

I) This idea, namely *conditioned Fourier expansion of the increments* can be applied to more general cases. It can be "random walks on a graph/group/Polish space" , "discrete-time Markov chains on a manifold", or "general semi-martingales", etc.

II) It is also notable that our DIF holds irrespective of the distributions, as far as the reference measures are equivalent.

III) To the best of the author's knowledge, discrete Itô's formula was pioneered by T. Szabados [8].

IV) Discrete stochastic calculus, which have more emphasis on chaos expansions, has been studied by many. A nice exposition [4] on this topic is available.

3 Stochastic Numerics from the Perspective of Discrete Itô Calculus

3.1 Reduction to finite difference schemes

When ν is concentrated on a finite set $G = \{g_1, ..., g_l\}$, then

$$u_T^N(t,x) = \mathbf{E}[f(\mathbf{X}_T)|X_t = x], \quad t = t_0, t_1, ..., t_N, \ (t_N = T, \Delta t \equiv 1/N)$$

solves a finite difference equation

$$(\partial_t^N + L^N)u = 0 \text{ with terminal condition } u_T(T,x) = f(x), \tag{3.1}$$

where

$$\partial_t^N u(t,x) = N\{u(t,x) - u(t-1/N,x)\}$$

and

$$L^N u(t,\mathbf{x}) = N\sum_{i=1}^{l}\{u(t,\mathbf{x}+F(\mathbf{x},1/N,\mathbf{H}(g_i))) - u(t,\mathbf{x})\}\,\nu(g_i).$$

This means that if $\partial_t^N + L^N$ is *consistent* [1] with a differential operator $\partial_t + L$; i.e.

$$\int_t^T (\partial_t^N + L^N)\varphi\, ds \to \int_t^T (\partial_t + L)\varphi\, ds \text{ as } N \to \infty \text{ for any smooth } \varphi, \tag{3.2}$$

u^N converges to a smooth solution u of $(\partial_t + L)u = 0$.

Following standard arguments in the context of the finite difference scheme, we illustrate what is going on here. For the solution v

$$(\partial_t^N + L^N)v = \Phi, \ v(T) = 0, \tag{3.3}$$

we have a discrete Feynman-Kac formula:

$$v(t_k,x) = \sum_{l=k}^{N} \mathbf{E}[\Phi(t_l,\mathbf{X}_{t_l})|\mathbf{X}_{t_k} = x].$$

[1] See, e.g. [6]

Suppose that the smooth solution u exists. Then $u - u^N$ is a solution to (3.3) with $\Phi = \{(\partial_t^N + L^N) - (\partial_t + L)\}u$. Therefore,

$$|u^N(t,x) - u(t,x)| \leq \sum |\mathbf{E}[\{(\partial_t^N + L^N) - (\partial_t + L)\}u(t_l, X_{t_l})|X_t = x]|,$$

and the consistency (3.2) gives the convergence.

Namely it acts as a finite difference approximation of the partial differential equation. Note that the problem of the rate of convergence in the Euler-Maruyama scheme:

$$\mathbf{E}[f(\mathbf{X}_t^N)] \to \mathbf{E}[f(\mathbf{X}_T)]$$

reduces to the same problem in (3.2), which can be easily calculated in many cases.

3.2 Completeness makes it slow

As we remarked in section 2.6, the cases where $\sharp G = 1+$ the dimension of the state space is of special interest since it serves as a complete market model. However, if one considers them to be a discretization, by the Euler-Maruyama scheme, of a continuous-time (complete market) model, one is obliged to pay some costs.

Theorem 3.1 ([1]). *The unique prices of European claims in discrete-time complete markets converge to the ones in the corresponding continuous complete market as the time-step Δt tends to 0. The order of convergence is at least $\sqrt{\Delta t}$ and* **it cannot be improved** *when $n \geq 2$.*

Roughly speaking, this is because the set G with $\sharp G = n + 1$ **cannot support any** n-dimensional random variable which has the same moments up to degree three with the increment of n-dimensional Brownian motion, when $n \geq 2$. For details, see [1].

3.3 "Infinite" difference scheme

The argument in section 3.1 is still valid for a general ν by putting

$$L^N f(\mathbf{x}) = N \int f(\mathbf{x} + F(\mathbf{x}, N^{-1}, \mathbf{H}(y)))\, \nu(dy) \tag{3.4}$$

for the SDE's in section 2.5.

We propose the following implementations.

1. Let ν be the Lebesgue measure on $[0, 1]$.

2. Let F be as (2.6) (Euler-Maruyama).

3. Take the Walsh system as an ONB. Here by the Walsh system we mean the group generated by

$$\tau_n(x) = \begin{cases} +1 & \text{if } 2k \leq 2^n x < 2k+1 \text{ for } k = 0, 1, ..., 2^{n-1} - 1, \\ -1 & \text{otherwise,} \end{cases}$$

$n = 1, 2, \cdots$. Note that $\tau_1, ..., \tau_n, ...$ are nothing but a Bernoulli sequence. We can take, for example,

$$\mathbf{H} = (\tau_1, \tau_2, \tau_3, \tau_1\tau_2\tau_3, \tau_4, \tau_5, \tau_1\tau_2\tau_4, \tau_1\tau_2\tau_5, \tau_1\tau_2\tau_3\tau_4\tau_5, ...).$$

(Just avoid using those with the even-number length.)

4. Simulate the path $\mathbf{X}$ by a Monte-Carlo/Quasi Monte-Carlo uniform sequence in $[0, 1)^N$.

This method is meant to be a Monte-Carlo simulation scheme of high-dimensional finite difference scheme. It is almost dimension-free. In fact, the dimension n can be effectively very large; around 3000, as is reported in [7].

3.4 Supplementary remarks for section 3

The correspndence of the general Markov chian approximation by (2.4) to finite-difference shemes can be generalized to non-linear cases. This generalization includes the Kushner's correspondence (see e.g. [5]).

References

[1] Akahori, J.: A discrete Ito calculus approach to He's framework for multi-factor discrete market. preprint.

[2] Ethier, S. N. and Kurtz, T. G. : *Markov processes. Characterization and convergence.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986.

[3] Fujita, T. :A random walk analogue of Lévy's theorem, preprint, 2003.

[4] Gzyl, H. : An exposé on discrete Wiener chaos expansions, preprint, January, 2005.

[5] Kushner, H. J. : *Probability methods for approximations in stochastic control and for elliptic equations.* Mathematics in Science and Engineering, Vol. 129. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1977.

[6] Richtmyer, R.D. and Morton, K.W. : *Difference methods for initial-value problems.* Second edition. Interscience Tracts in Pure and Applied Mathematics, No. 4 Interscience Publishers John Wiley & Sons, Inc., New York-London-Sydney (1967) [Reprint of the second edition. Robert E. Krieger Publishing Co., Inc., Malabar, FL, 1994]

[7] Sawai, T.: Simulations of SDE in very high dimensions. Master Thesis, Ritsumeikan University, 2006.

[8] Szabados, T. :A discrete Ito's formula, *Limit theorems in probability and statistics, Colloq. Math. Soc. Janos Bolyai* **57** (1989) 491–502.